

# The Gap Between Simulation and Understanding in Climate Modeling

Isaac M. Held

Geophysical Fluid Dynamics Laboratory/NOAA

Princeton, NJ

Submitted to the *Bulletin of the American Meteorological Society*

September 11, 2004

## **Abstract**

The problem of creating truly convincing numerical simulations of our Earth's climate will remain a challenge for the next generation of climate scientists. Hopefully, the ever-increasing power of computers will make this task somewhat less frustrating than it is at present. But increasing computational power also raises issues as to how we would like to see climate modeling and the study of climate dynamics evolve in the 21st century. One of the key issues we will need to address is the widening gap between *simulation* and *understanding*.

# 1 The need for model hierarchies

The complexity of the climate system presents a challenge to climate theory, and to the manner in which theory and observations interact, eliciting a range of responses. On the one hand, we try to *simulate* by capturing as much of the dynamics as we can in comprehensive numerical models. On the other hand, we try to *understand* by simplifying and capturing the essence of a phenomenon in idealized models, or even with qualitative pictures. As our comprehensive models improve in quality, they more and more often become the primary tools by which theory confronts observations. The study of global warming is an especially good example of this trend. A handful of major modeling centers around the world compete in creating the most convincing climate simulations and the most reliable forecasts of climate change, while large observational efforts are mounted with the stated goal of improving these comprehensive models.

Due to the great practical value of simulations, and the opportunities provided by the continuing increases in computational power, the importance of *understanding* is occasionally questioned. What does it mean, after all, to understand a system as complex as the climate, when we cannot fully understand idealized nonlinear systems with only a few degrees of freedom?

Without attempting an all-encompassing definition, it is fair to say that we typically gain some understanding of a complex system by relating its behavior to that of other, especially simpler, systems. For sufficiently complex systems, we need a model hierarchy on which to base our understanding, describing how the dynamics changes as key sources of complexity are added or subtracted. Our understanding benefits from appreciation of the interrelationships among all elements of the hierarchy.

The importance of such a hierarchy for climate modeling and studies of atmospheric and oceanic dynamics has often been emphasized. See, for example, Schneider and Dickinson (1974) and especially Hoskins (1983). But despite notable exceptions in a few subfields, research on ENSO being one example, climate theory has not, in my opinion, been very successful at hierarchy construction. By this statement I do not mean to imply that important work has not been performed, of course, but only that the gap between comprehensive climate models and more idealized models has not been successfully closed.

Consider by analogy another field that must deal with exceedingly complex systems – molecular biology. How is it that biologists have made such dramatic and steady progress in sorting out the human genome and the actions and interactions of the thousands of proteins of which we are constructed? Without doubt, one key has been that nature has provided us with a hierarchy of biological systems of increasing complexity amenable to experimental manipulation, ranging from bacteria to fruit fly to mouse to man. Furthermore, the nature of evolution assures us that much of what we learn from simpler organisms is directly relevant to deciphering the workings of their more complex relatives. What good fortune for the biological sciences to be presented with precisely the kind of hierarchy needed to understand a complex system! Imagine how much progress would have been made if one were limited to studying man alone.

Unfortunately, Nature has not provided us with simpler climate systems that form such a beautiful hierarchy. Planetary atmospheres provide us with some insights into the range of behaviors possible, but they are few in number, and each planet has its own idiosyncrasies. While their study has connected to terrestrial climate theory/modeling on occasion, the influence

has not been systematic. Laboratory simulations of rotating and/or convecting fluids remain a valuable and underutilized resource, but they cannot address many of our most complex problems. We are left with the necessity of constructing our own hierarchies of climate models.

Because nature has provided the biological hierarchy, it is much easier to focus the attention of biologists on a few representatives of the key evolutionary steps towards greater complexity. And such a focus is central to success. If every molecular biologist had simply studied his or her own favorite bacterium or insect, rather than focusing so intensively on *E. coli* or *Drosophila melanogaster*, it is safe to assume that progress would have been far less rapid.

It is emblematic of our problem that studying the biological hierarchy is *experimental* science, while constructing and studying climate hierarchies is *theoretical* science. One can justify studying *E. coli*. not only because it shares many fundamental genetic mechanisms with all cells, but also because it exists, after all, and it and its close bacterial relatives affect the world in ways that are worth understanding at the molecular level in their own right. Elements of a climate model hierarchy are generally only of interest to climate theorists.

A biologist need not convince her colleagues that the model system she is advocating for intensive study is well-designed or well-posed, but only that it fills an important niche in the hierarchy of complexity and that it is convenient for study. Climate theorists are faced with the difficult task of both constructing a hierarchy of models and somehow focusing the attention of the community on a few of these models so that our efforts accumulate efficiently. Even if one believes that one has defined the *E. coli* of climate models, it is difficult to energize (and fund) a significant number

of researchers to take this model seriously and devote substantial parts of their careers to its study.

And yet, despite the extra burden of trying to create a consensus as to what the appropriate climate model hierarchies are, the construction of such hierarchies must, I believe, be a central goal of climate theory in the 21st century. There are no alternatives if we want to understand the climate system and our comprehensive climate models. Our understanding will be embedded within these hierarchies.

## **2 The practical importance of understanding**

Why should we care that we do not understand our comprehensive climate models as dynamical systems in their own right? Does this matter if our primary goal happens to be to improve our simulations, rather than to create a subjective feeling of satisfaction in the mind of some climate theorist?

Suppose that one can divide a climate model into many small distinct components and that one can devise a testing and development strategy for each of these modules in isolation (including the form of the interactions among these modules). If the components have been adequately tested, is there any need for an understanding of what happens when they are coupled? To the extent that one can break down the testing process into manageable pieces, this bottom-up, reductive strategy is without doubt an appropriate and efficient approach to model development. Understanding is needed at the level of the module in question, so as to ensure its fidelity to nature, but is there understanding to be gained as a higher, more holistic level, that is of value to the climate modeling enterprise? Are we better off limiting ourselves to trying to understand particular physical processes of climatic

relevance?

The radiation code in atmospheric models (the clear-sky component, at least) is a good example. The broad band computations used in climate models are systematically tested against line-by-line computations based on the latest laboratory studies and field programs. When atmospheric observations and/or laboratory absorption studies require a modification to the underlying data base (for example, with regard to water vapor continuum absorption) this new information makes its way more or less efficiently into the broad-band climate model codes. Given this relatively convincing methodology, the clear-sky radiative flux component of climate models is generally treated with respect, evolving only when driven to do so by evidence of the sort outlined above.

Work towards devising similar methodologies for other model components is obviously of vital importance. But we are very far today from being able to construct our comprehensive climate models in this systematic fashion. Despite several major observational campaigns designed to guide us towards appropriate closures for deep moist convection, as an important example, there is little sense of convergence among existing atmospheric models. A program in which cloud-resolving simulations are systematically used as a middle ground between closure schemes and observations promises to improve this situation in the future, but there is still a long way to go.

When a fully satisfactory systematic bottom-up approach to model building is unavailable, the development process can be described, without any pejorative connotations intended whatsoever, as engineering, or even tinkering. (Our most famous inventors are often described as tinkerers!) Various ideas are put forward by the team building the model, based on their wisdom and experience, as well as their idiosyncratic interests and prejudices. To

the extent that a modification to the model based on these ideas helps ameliorate a significant model deficiency, even if it is, serendipitously, a different deficiency than the one providing the original motivation, it is accepted into the model. Generated by these informed random walks, and being evaluated with different criteria of merit, the comprehensive climate models developed by various groups around the world evolve along distinct paths.

The value of a holistic understanding of climate dynamics for model development is in making this process more informed and less random, and thereby more efficient. To the extent that we have little understanding of which aspects of a moist convection scheme are most important for exaggerating the double ITCZ in the East Pacific, or which help control the period of ENSO, then our search for ways to ameliorate our double ITCZ or improve our ENSO spectrum will be that much more random and less informed.

A holistic understanding of climate dynamics also helps in relating one comprehensive model to another. To the extent that we have some understanding of which aspects of convection schemes, or boundary layer models, or models of momentum exchange with the surface, matter for various aspects of our climate simulations, we can appreciate why one simulation is better than another without systematically and laboriously morphing one model into the other. One should then be able to take advantage more efficiently of the successes of other models.

As another example, if stratosphere-troposphere interactions in one comprehensive model result in a trend in the North Atlantic Oscillation as a result of increasing carbon dioxide, but not in other models, how does one judge which is correct? One can try to judge which model has the most realistic stratospheric-tropospheric interactions by comparing against ob-



servations, with the understanding that theoretical guidance is required to design these comparisons. One can also analyze more idealized models designed to capture the essence of the interaction in simpler contexts, within which the climate dynamics community can focus directly on the central issues. These idealized studies can then suggest optimal ways of categorizing or analyzing more comprehensive models. If we are to claim some understanding of this issue, our modeling hierarchy, from idealized to realistic, should tell a consistent story.

The climate simulation community does organize itself to perform a large variety of CMIPs (Climate Modeling Intercomparison Projects), those underlying the IPCC (Intergovernmental Panel on Climate Change) being the best known. By avoiding unnecessary differences between calculations, and getting different groups to compare results carefully, these projects have demonstrated their value, especially in teaching us which aspects of simulations are robust and which are not. CMIPs that involve integrating the models with idealized boundary conditions (for example, the Aqua-Planet Experiment Project, <http://www-pcmdi.llnl.gov/amip/ape>) potentially have an important role to play in relating high-end simulations and more idealized models. But CMIPs are not enough, as researchers involved in them are well aware. Comparisons of complex models that are not easily morphed into each other, even if forced with idealized boundary conditions, invariably leave us without a satisfactory understanding of why models differ. One must simplify the models as well as the boundary conditions.

### 3 Constructing a hierarchy

Even the simplest levels of the hierarchies that I have in mind are turbulent and chaotic models that one cannot hope to understand in all detail. This is not meant to imply that even simpler models do not have important roles to play. The simpler the model that explains some aspect of climate dynamics the better! But the claim is that there are sources of complexity in the climate system that prevent us from generating convincing simple quantitative theories for many of the questions that interest us. My concern here is with models that attack some of the core sources of complexity in the climate system, that allow one to address questions of climate maintenance and sensitivity, and that cannot be fully solved by an individual researcher but rather require the concerted efforts of a variety of investigators if meaningful progress is to be made.

I list a few examples with which I am particularly familiar, meant to be illustrative only. It is not the job of an individual or a small committee to decide what the appropriate model hierarchies are; rather models must prove themselves over time, and as they do so hierarchies should emerge naturally. Yet, given the relatively small number of researchers in climate theory, we need to make this process as efficient as possible, a point to which I will return below. The examples are all atmospheric.

A traditional choice for our *E. Coli* might be Phillips' (1956) original general circulation model, or, as we would refer to it today, the two-layer quasi-geostrophic (QG) model of a statistically steady baroclinically unstable jet on a  $\beta$ -plane, forced by linear radiative damping and linear surface friction. A snapshot of the upper layer potential vorticity from such a model is shown in Figure 1.

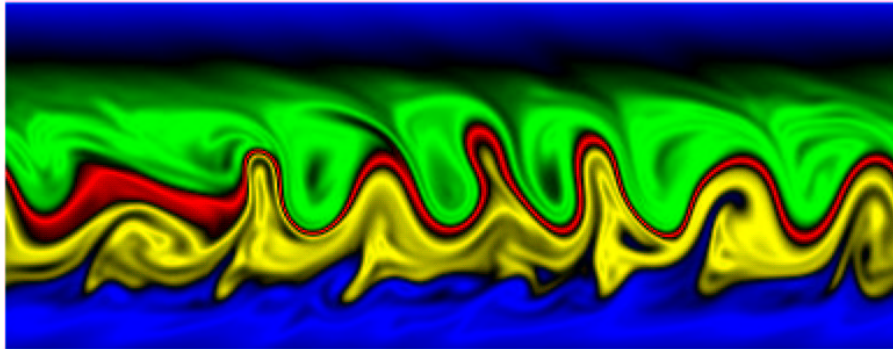


Figure 1: Upper layer potential vorticity in a statistically steady, zonally homogeneous, baroclinically unstable jet on a  $\beta$ -plane

This simple dynamical system generates flows with a rich structure that, many of us feel, captures the essence of midlatitude storm track dynamics. Despite efforts over the years, there is still much to be understood concerning the spectrum of the eddies, the coherence of the wave packets into which the baroclinic waves are typically organized (Figure 2 is a snapshot of a coherent wavepacket), and the maintenance of the time mean jet structure, as a function of the model parameters. This model also comes into its fair share of criticism, mostly related to the way in which it distorts linear baroclinic instability. My intuitive impression is that it is, in fact, a better model of the finite amplitude statistically steady storm track than it is of weakly nonlinear disturbances. It is in the nature of hierarchy development that the simpler members will be missing essential ingredients of the real system. The decision we are constantly making is whether, despite these deficiencies, what is retained is still of importance. The analysis of this classical system has gone out of fashion; as a result, we try to discuss the behavior of storm tracks in more complex models without a consensus on

their behavior in this simplest of dynamical frameworks.

Taking several steps towards greater complexity, one can consider dry primitive equation models on the sphere, forced in the simplest possible ways once again, for example by linear thermal relaxation and linear surface friction. One version of such a model was described by Held and Suarez(1994). Being dry, the focus is still on midlatitude dynamics, but a variety of new issues arise because of the spherical geometry and the non-QG dynamics. The former allows one to meaningfully analyze factors that control the structure of the eddy momentum fluxes and the distribution of surface westerlies, the latter results in more realistic frontal dynamics, and the combination results in more realistic eddy life cycles. Additional numerical issues, related to gravity wave generation and convergence also arise in this context. Fig. 2 is a snapshot of the lower tropospheric vertical motion and temperature fields on the sphere from such a model.

Bridging the gap between this kind of model and comprehensive atmospheric climate models requires the development of idealized dynamical frameworks in which one can discuss and analyze the *moist general circulation*, including the role of moist convection, and the simulation of the Earth's cloud distribution. I think it is fair to say that it is in models of the global moist general circulation that many of us feel most strongly the presence of a gap between simulation and understanding.

Fig 3 shows some results from a model being constructed by Dargan Frierson and collaborators at Princeton University, which we hope will be relevant in this regard. It solves for the flow of an ideal gas on a rotating sphere, contains very simple grey radiative transfer that is a function only of temperature, a highly simplified boundary layer mixing scheme, a series of different convective closures which are the focal point of this study, but

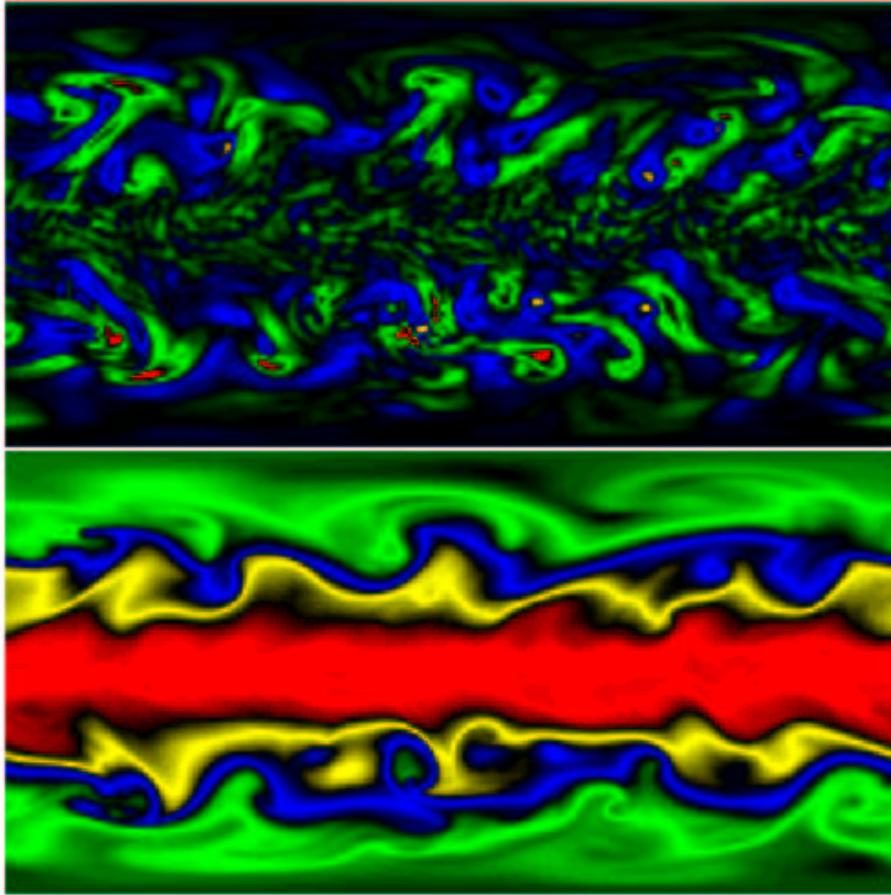


Figure 2: Snapshot of vertical motion (upper panel) and lower tropospheric temperature (lower panel) for the dry spherical model of Held and Suarez (1994)

no condensate, and a homogeneous surface. The choice here has been to use a slab ocean ("mixed layer") boundary condition, rather than fixed surface temperatures, so as to facilitate the discussion of climate sensitivity. Moist processes complicate climate models through the effects of latent heat release on atmospheric circulations, and through the effects that the vapor and condensate have on radiative heating. Here we ignore the latter to try to isolate the direct effects of latent heat release. The different panels show snapshots of the rate of precipitation using three distinctly different convection schemes. We need to understand which aspects of convective closures control the disparate structures of the Intertropical Convergence Zone seen in these three panels. We also need to insure that we can generate results such as these in a robust and reproducible manner. (There may very well be details of implementation that we tend to ignore in documenting our models that affect these solutions.)

This particular dynamical framework is unlikely to be ideal, but I am convinced that clearly defined idealized models of this sort are urgently needed. We especially need careful studies of cloud feedbacks in controlled and reproducible model settings to supplement studies of climate sensitivity with comprehensive models. We also need clean results on how simulated tropical storm climatologies depend on the model formulation. Details of numerical implementation, convergence and sensitivity to resolution becomes especially important in these contexts.

In each of the three examples described above, the models produce climates that are independent of longitude, but each can be modified to create zonal inhomogeneities. There have been bursts of activity in the past involving idealized models of the zonally asymmetric climatic response to orography and land-ocean geometry. We need another sustained burst, taking

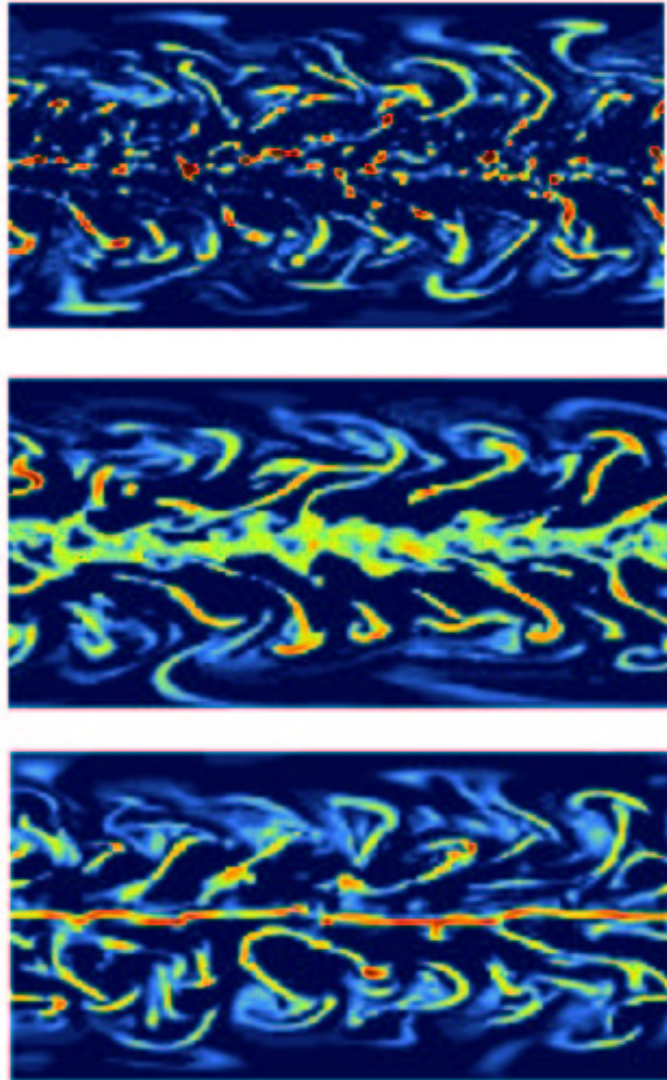


Figure 3: Instantaneous precipitation over the globe from three version of an idealized moist general circulation model, using three convective closures. Courtesy of D. Frierson

advantage of increases in computational power, building on past progress and on careful studies of models with zonally symmetric climates, focusing in part on how the storm tracks and stationary waves mutually shape each other, as well as on the interactions between land surfaces and the hydrological cycle, hopefully contributing to a firmer foundation for the analysis of regional climate change. Monsoon dynamics remains an especially exciting area of focus in this regard.

We would all prefer to model tropical convection at much higher resolution to remove some of the dependence of our solutions on convective parameterizations. An alternative starting point for a modeling hierarchy that is attracting growing interest is that of non-rotating radiative-convective equilibrium in a moist atmosphere. In a domain that is homogeneous in the horizontal, one studies the moist convective turbulence that is generated by uniform solar heating, over a uniform fixed-temperature saturated surface or a saturated surface with a specified heat capacity. One can perform this calculation with a global model, with parameterized deep convection. But more typically, one pushes to much higher horizontal resolution, in a doubly periodic domain in which one ignores spherical geometry. Today these models typically have horizontal resolution ranging from 1 to 4 km, so as to explicitly resolve at a minimum the largest convective cores that extend to the tropopause (see Figure 4), but there is desire to proceed to even finer resolution so as to explicitly resolve shallow convection, the associated low level cloud field, and ideally the energy-containing eddies in the boundary layer that shape these fields. In this relatively simple setting one can test one's understanding of the maintenance of the tropical humidity distribution, the amount of conditional instability in the equilibrated state, the intermittency of the convection, and the mean liquid water and ice concentrations. This



model can be pushed in many interesting directions, incorporating rotation for example.

Our best attempts at constructing comprehensive cloud-resolving models obviously revolve around real case studies and careful comparisons with all available data. But from the perspective of theory, just as for global models it will not be sufficient to integrate our most comprehensive convection-microphysical-radiative model in idealized and/or realistic configurations and see what we get. We will also need models in idealized geometries in which we vary (and simplify) the assumptions concerning the microphysics of hydrometeors and cloud-radiative interactions. Ideally, we will eventually be able to come to agreement on how the planetary albedo, for example, depends on the microphysical/radiative formulation. We also need to compare aspects of these predictions to specific coarse resolution models using different classes of convective closures, so as to make contact with existing high-end simulations. We will need a hierarchy of models of homogeneous radiative-convective equilibrium. And we will need to build bridges between this homogeneous framework and inhomogeneous systems of interest.

Whenever we work in an idealized theoretical framework such as this, we are always worrying that we are not studying the right system. Is homogeneous moist radiative-convective equilibrium in the absence of environmental shear or rotation too unrealistic to be useful? Should one move immediately towards simple inhomogeneous environments, such as a non-rotating Walker cell? If we could create a simulation of moist radiative convective equilibrium in the laboratory, much like Benard convection, we would be closer to the position of a molecular biologist studying the development of the fruit fly: we would have a far clearer picture of the appropriateness of the system that we are studying. Unfortunately, we have no such laboratory simulation.

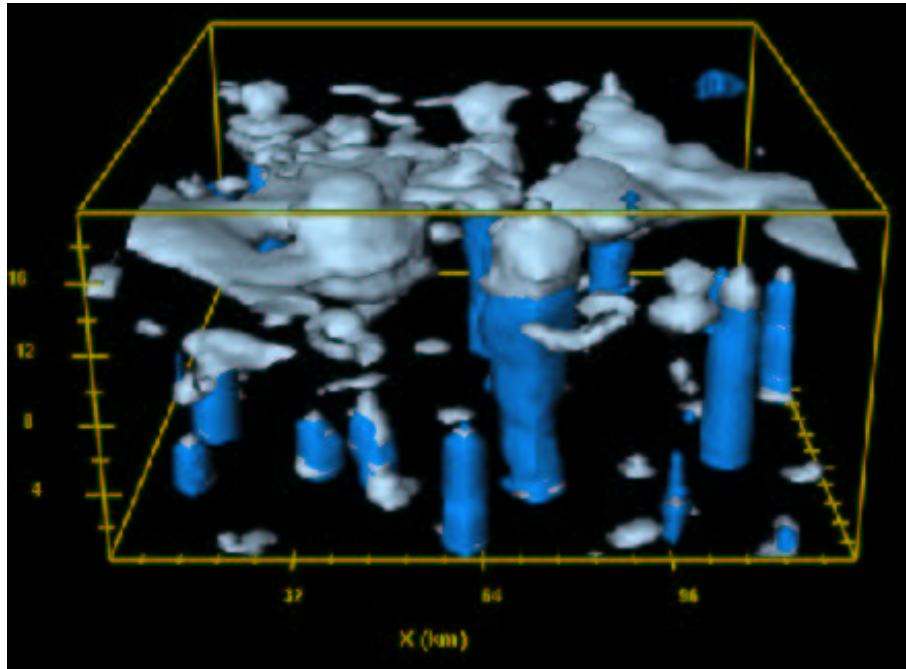


Figure 4: Radiative-convective equilibrium over a 128 km square domain extending into the lower stratosphere, with falling condensate colored blue and suspended condensate colored grey

Think of how revolutionary it would be for studies of moist convection if we did!

That we have such a variety of potentially useful idealized models, even within the atmosphere in isolation from the oceans, ice, and land surface, serves to emphasize the difficulty of our task.

## 4 The future of climate theory

Accepting that the kind of understanding that emerges from the construction and analysis of climate model hierarchies is important, and given the many efforts underway that are devoted to models of various levels of complexity, are there things we could do to make this effort more productive? I highlight two related tendencies which have slowed the systematic development of climate model hierarchies. (My own work illustrates these tendencies nicely, and this discussion is as much a self-critique as it is one of the field more generally.)

### 4.1 Elegance vs. elaboration

Our goal must be to reduce the number of idealized models that we analyze. Otherwise we are left with a string of more or less interesting results, few of which have been intensively examined by more than 2 or 3 people, and which we never quite manage to relate to each other. Furthermore, when models are underanalyzed and not reproduced by others, we are never certain that the computations have been performed properly. But how can this inefficient deployment of our theoretical resources be avoided? The key, I feel, is *elegance*.

An elegant model is as elaborate as it needs to be to capture the essence

of a particular source of complexity, but no more elaborate. Many of our models are more elaborate than they need be, and this is, I believe, the prime reason why it is difficult for the field as a whole to focus efficiently on a small number of models. If a particular scheme seems unnecessarily baroque, why should I use it as a basis for my own research? What lasting value will my study have?

Over-elaboration is the (understandable) result of the pressure we all feel for our work to be relevant to the big issues in climate dynamics. This relevance generally requires a certain level of realism in one's simulations, and this pressure to reach the required level of realism pushes models towards ever-increasing elaboration. Yet, in the process one's model often loses much of its attraction to other researchers, who may not be in agreement with all of the choices made in the process of elaboration.

We justify our research, to ourselves and others, by appealing to some mixture of short-term practical consequences and lasting value. High end simulations are primarily driven by the need to meet practical applications, requiring them to be as realistic as possible given existing resources. These simulations need be of no lasting value, as they will be supplanted by ever more comprehensive models as computer resources increase. When global nonhydrostatic atmospheric models resolving deep moist convection become common in future decades, the global warming simulations obtained with the current generation of models will be of historical interest only. But the importance of the problem is such that we cannot wait for this to occur; we need to do our best now, knowing full well that these efforts will be obsolete within most of our lifetimes. While there is no value in elaborating these comprehensive simulations in ways that have no practical consequences or no hope of confronting data, an emphasis on elegance can be counterproductive;

a large number of details may very well be needed to get a useful simulation.

As we back off from this high end, the balance between elegance and realism becomes more of an issue. My reading of the literature is that elegance is often sacrificed unnecessarily, partly for the sake of a competition with comprehensive models. The latter seem, after all, to be extraordinarily inefficient at attacking many key climate problems. Yet, in an era of exponentially increasing computational power, this competition is often less valuable than we might like to admit, given the time scale at which studies become feasible at a more comprehensive level.

It may very well be that we need fewer idealized climate models, but that we need a *larger* number of comprehensive models! Given the large number of choices that must be made in the construction of a comprehensive climate model, and the complexity of the fitness function that one is trying to optimize, there is clearly value in trying to sample more widely in the space of possible models. Given the difficulty of creating a single model, both in human and computational resources, this seems paradoxical, but the efforts at the Hadley Centre at creating an ensemble of climate models (Murphy, et. al., 2004) are encouraging in this regard. Favoring a large number of such models is consistent with the claim that a given comprehensive model is not constructed with lasting value as the primary goal.

Elegance and lasting value are correlated. An elegant hierarchy of models upon which the field as a whole bases its understanding of the climate system can be of benefit to future generations for whom our comprehensive simulations, valuable as they are at present, will have become obsolete.

## 4.2 Conceptual research vs. hierarchy development

A theoretically inclined researcher might design and build a model, on the basis of which he or she tries to create the case for some picture of a phenomenon or for the utility of some concept or approach – and then discard the model. The model is not intended, in many cases, to have a life of its own, but is rather a temporary expedient. In the limiting case, the model is not fully described and the result not fully reproducible. Or an existing model might be used in the same way. The focus is on the concept being put forward, not on the model itself. I refer to this as *conceptual research*. Much of the best work with comprehensive models can be classified as conceptual, as can, for example, much of the paleoclimatic research with computationally efficient climate models of *intermediate complexity*, as they are often called. In this context the model is a useful tool that helps one think about the system and search for ways in which to interpret observations.

Some might argue that all modeling is conceptual in this sense, that all models are just expedient tools and not themselves the final goal, and that individual models never deserve to be thought of as having lasting value. Given the level of complexity that we face in the climate problem, I do not think that this is a viable perspective. As I have tried to argue throughout this essay, without the solid foundation provided by careful study of an appropriate model hierarchy, there is a danger that we will be faced with a babel of modeling results that we cannot in any satisfying way relate to one another. Our modeling activity must have as its goal the creation of models of lasting value, in addition to facilitating conceptual research. Ideally, we need some models of intermediate complexity that we take just as seriously as do the biologists who map out every single connection in the nervous

system of the snail!

## 5 Concluding Remarks

The health of climate theory/modeling in the coming decades is threatened by a growing gap between high-end simulations and idealized theoretical work. In order to fill this gap research with a hierarchy of models is needed. But to be successful, this work must make progress towards two goals simultaneously. It must, on the one hand, make contact with the high end simulations and improve the comprehensive model development process; otherwise it is irrelevant to that process, and, therefore, to all of the important applications built on our ability to simulate. On the other hand, it must proceed more systematically towards the creation of a hierarchy of lasting value, providing a solid framework within which our understanding of the climate system, and that of future generations, is embedded. Funding for climate dynamics should reflect this need to balance conceptual research, simulation, and hierarchy development.

## REFERENCES

- Held, I. M., and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Am. Meteor. Soc.*, **75**, 1825-1830.
- Hoskins, B. J., 1983: Dynamical processes in the atmosphere and the use of models. *Quart. J. Roy. Met. Soc.*, **109**, 1-21.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**, 768 - 772 (12 August).
- Phillips, N. A., 1956: The general circulation of the atmosphere: a numerical experiment." *Quart. J. Roy. Meteor. Soc.*, **82**, 123-164.
- Schneider, S. H., and R. E. Dickinson, 1974: Climate Modeling, *Reviews of Geophysics and Space Physics*, **12**, 447 - 493.